

Rao-Blackwellised Interacting Markov Chain Monte Carlo for Electromagnetic Scattering Inversion

F. Giraud¹, P. Minvielle¹, M. Sancandi¹, P. Del Moral²,

¹ CEA-CESTA, 33114 Le Barp, France

² INRIA Bordeaux Sud-Ouest, Domaine Universitaire, 351, cours de la Liberation, 33405 Talence Cedex, France

E-mail: francois.giraud@cea.fr

Abstract. The following electromagnetism (EM) inverse problem is addressed. It consists in estimating local radioelectric properties of materials recovering an object from the global EM scattering measurement, at various incidences and wave frequencies. This large scale ill-posed inverse problem is explored by an intensive exploitation of an efficient 2D Maxwell solver, distributed on High Performance Computing (HPC) machines. Applied to a large training data set, a statistical analysis reduces the problem to a simpler probabilistic metamodel, on which Bayesian inference can be performed. Considering the radioelectric properties as a dynamic stochastic process, evolving in function of the frequency, it is shown how advanced Markov Chain Monte Carlo methods, called Sequential Monte Carlo (SMC) or interacting particles, can provide estimations of the EM properties of each material, and their associated uncertainties.

1. Introduction

The inverse problem is described in figure 1. The Radar Cross Section quantifies the scattering power of an object, at a given incidence and wave frequency. It is defined as the ratio between the radar transmitted power and the incident power density (in plane wave) [1]. RCS measurement process is schematically presented on the right part of figure 1. The object or mock-up is illuminated by a quasi-planar monochromatic wave, inside an anechoic chamber where interferences are limited. The acquisitions are realized at K successive discrete frequencies (f_1, f_2, \dots, f_K), for different incidence angles (by piloting motorized rotating support). From this raw data, a signal processing is performed, mainly consisting of calibration and filtering. At the end, the measurement provides an evaluation of the calibrated complex (amplitude and phase) scattering coefficient, for each frequency and incidence.

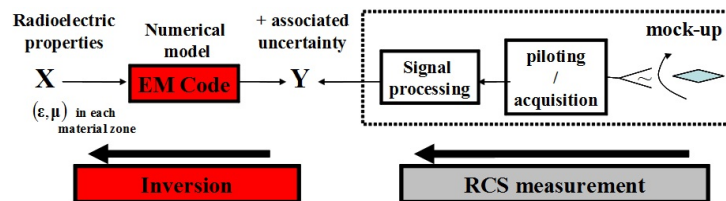


Figure 1. The EM inverse problem

A metallic axi-symmetric object is recovered with N areas (see figure 3), each area corresponding to a material with its associated isotropic radioelectric properties, i.e. the complex parameters of permittivity ϵ and permeability μ . The EM inverse problem can be expressed as: is it possible to extract some local information on the material properties (ϵ_k, μ_k) of each area from the global scattering measurement?

2. The stochastic modeling

At a given frequency f_k , the system state \mathbf{X}_k can be defined by (omitting f_k to lighten the notations): $\mathbf{X}_k = [\underline{\epsilon}' \ \underline{\epsilon}'' \ \underline{\mu}' \ \underline{\mu}'']^T$, with $\underline{\epsilon}'$, $\underline{\epsilon}''$, $\underline{\mu}'$ and $\underline{\mu}''$, respectively the real and imaginary permittivity and permeability components of the N areas. Notice that the state dimension can be high ($4 \cdot N$ with $N \approx 100$). On the other hand, considering measurements at the given frequency f_k with various incidence angles in both polarizations (TM and TE), the observation vector is made of the real ($\Re(\cdot)$) and imaginary ($\Im(\cdot)$) parts of the complex scattering coefficients \mathbf{c}_{TM} and \mathbf{c}_{TE} measured at M angles $\theta_1, \dots, \theta_M$: $\mathbf{Y}_k = [\Re(\mathbf{c}_{\text{TM}}) \ \Im(\mathbf{c}_{\text{TE}}) \ \Re(\mathbf{c}_{\text{TM}}) \ \Im(\mathbf{c}_{\text{TE}})]^T$. Considering all frequencies, vectors $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_K)$ and $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_K)$ respectively define complete system state and observation.

2.1. Observation model

The 2D-axisymmetric Maxwell solver software (\mathcal{F}_{2D}) can predict the observation from the system state. Assuming a multidimensional Gaussian measurement uncertainty model, it leads to the following likelihood model $\mathbf{Y}_k | \mathbf{X}_k \sim \mathcal{N}(\mathcal{F}_{2D}(\mathbf{X}_k), \mathbf{R}_k)$ (at a given frequency f_k), where \mathbf{R}_k is the covariance matrix (assumed known). To avoid numerous and heavy \mathcal{F}_{2D} computations, we have developed the following global approach. First, the high dimension state space and the associated system response are explored randomly around expected properties (prior knowledge), computations being massively distributed on HPC machines.

Let $\mathcal{B}_A = \{(\mathbf{X}_k^{(1)}, \mathbf{Y}_k^{(1)}), \dots, (\mathbf{X}_k^{(N_E)}, \mathbf{Y}_k^{(N_E)})\}$ the training data composed of N_E couples. It is then processed by N-D statistical techniques; sensibility analysis and model reduction techniques can possibly reduce the state space dimension. Applying multidimensional regression, it turns out that the model is approximatively linear. According to the studied cases, the linearity errors, evaluated by residue analysis and bootstrap techniques, are significant but much lesser than the RCS measurement uncertainties. Finally, the following linear Gaussian model can be considered (at a given frequency f_k):

$$\mathbf{Y}_k | \mathbf{X}_k \sim \mathcal{N}(\mathbf{A}_k \cdot \mathbf{X}_k + \mathbf{Y}_k^0, \mathbf{R}_k) \quad (1)$$

where the deterministic part of the linear model is given by the learned matrix $4M \times 4N$ \mathbf{A}_k and the vector \mathbf{Y}_k^0 .

2.2. Prior model at a fixed frequency

Let us fix a frequency f_k . We model our a priori knowledge on \mathbf{X}_k with a Gaussian distribution $\mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$. The object is divided in N_b blocks of areas, each of them composed of a rather homogeneous material. The location of these blocks is known exactly. Prior mean value \mathbf{m}_k is defined with reference values of ϵ' , ϵ'' , μ' and μ'' for each of these blocks. Then, for any component \mathbf{X}_k^i , we define a variance $\sigma_{S,k}^i$ as a mix between absolute and relative uncertainty. To take into account the spatial local homogeneity, covariance matrix \mathbf{P}_k is defined block by block independently (and separately for each of the ϵ' , ϵ'' , μ' , μ'') by correlation relations between \mathbf{X}_k 's block-sharing components:

$$Cov(\mathbf{X}_k^i, \mathbf{X}_k^j) = \sigma_{k,S}^i \cdot \rho_S^{|i-j|}$$

where $\rho_S \in [0, 1]$ is a spatial correlation parameter (typically $\rho_S = 0.95$ in our applications). It means that the correlation between 2 areas decreases geometrically at speed ρ_S with the distance between them.

2.3. Inter-frequential prior model

Radioelectric properties ϵ' , ϵ'' , μ' , μ'' are known to vary in function of the wave frequency [1]; their non-stationary dynamic can be quite different from frequency f_1 to f_K . However, in order to take account of expected frequency profiles regularity (for each material and ϵ' , ϵ'' , μ' , μ''), we model sequence $(\mathbf{X}_k, k \in \{1, \dots, K\})$ with a linear Gaussian correlation structure, given by the generalised autoregressive (AR) random process [2]:

$$\begin{aligned} \mathbf{X}_1 &\sim \mathcal{N}(\mathbf{m}_1, \mathbf{P}_1) \\ (\mathbf{X}_{k+1} - \mathbf{m}_{k+1}) &= \mathbf{D}_\rho \cdot \mathbf{H}_{k+1} \cdot \mathbf{H}_k^{-1} \cdot (\mathbf{X}_k - \mathbf{m}_k) + \sqrt{\mathbf{I}_d - \mathbf{D}_\rho^2} \cdot \mathbf{H}_{k+1} \cdot \mathbf{V}_k \end{aligned} \quad (2)$$

where

- for each k , \mathbf{H}_k is the “square root” matrix of covariance matrix \mathbf{P}_k , defined as being the unique symmetric definite positive matrix satisfying $\mathbf{H}_k \cdot \mathbf{H}_k^T = \mathbf{P}_k$;
- $(\mathbf{V}_k, k \in \{1, \dots, K\})$ are i.i.d. $\mathcal{N}(0, \mathbf{I}_d)$;
- \mathbf{D}_ρ is a positive diagonal matrix commuting with the \mathbf{H}_k , further described, and depending on a parameter ρ .

One notices that so defined sequence \mathbf{X}_k still admits $\mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$ as marginal distributions. More generally, it can be shown that the distribution of concatenated vector $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_K)$ is a Gaussian distribution with mean $\mathbf{m} = (\mathbf{m}_1, \dots, \mathbf{m}_K)$ and covariance matrix:

$$\mathbf{P} = \mathcal{H} \cdot \begin{pmatrix} \mathbf{I}_d & \mathbf{D}_\rho & \mathbf{D}_\rho^2 & \dots & \mathbf{D}_\rho^{K-1} \\ \mathbf{D}_\rho & \mathbf{I}_d & \mathbf{D}_\rho & & \vdots \\ \mathbf{D}_\rho^2 & \mathbf{D}_\rho & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbf{D}_\rho \\ \mathbf{D}_\rho^{K-1} & \dots & \dots & \mathbf{D}_\rho & \mathbf{I}_d \end{pmatrix} \cdot \mathcal{H}^T$$

where \mathcal{H} is the diagonal by block matrix $\mathcal{H} = \text{diag}(\mathbf{H}_1, \dots, \mathbf{H}_K)$, which explicits every joint distribution $(\mathbf{X}_i, \mathbf{X}_j)$.

Finally, let us clarify parameter ρ (and matrix \mathbf{D}_ρ). \mathbf{D}_ρ 's role in these equations makes clear that it's a frequential correlation parameter. Its dimension is to be chosen among 3 possibilities according to our assumptions.

- 1st case: ρ is assumed independant from material and ϵ' , ϵ'' , μ' , μ'' . It is one-dimensional, $\rho \in [0, 1]$, and $\mathbf{D}_\rho = \rho \cdot \mathbf{I}_d$.
- 2nd case: ρ depends on the material (block). It is N_b -dimensional, $\rho \in [0, 1]^{N_b}$, and \mathbf{D}_ρ giving to each line (component of \mathbf{X}_k) its associated component of ρ .
- 3rd case: ρ depends on material and ϵ' , ϵ'' , μ' , μ'' . It is $4 \cdot N_b$ -dimensional, and \mathbf{D}_ρ is the diagonal matrix composed with $4 \cdot N_b$ $\rho_i \cdot \mathbf{I}_d$ -type blocks.

2.4. Global model

Conditionnaly to frequential correlation parameter ρ , the problem of the determination of $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_K)$ given the measurements $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_K)$ can be expressed as a classic linear Gaussian hidden dynamic Markov process observed at "times" f_k ($k = 1, \dots, K$):

$$\mathbf{X}_{k+1} = \mathbf{M}_k^\rho \cdot \mathbf{X}_k + \mathbf{w}_k^{(1)} \quad \text{and} \quad \mathbf{Y}_k = \mathbf{A}_k \cdot \mathbf{X}_k + \mathbf{w}_k^{(2)}$$

where $\mathbf{w}_k^{(1)}$ and $\mathbf{w}_k^{(2)}$ are independant Gaussian noises with known parameters, and \mathbf{M}_k^ρ a known matrix depending on ρ (see (1) and (2)). Parameter ρ , intuitively representing the inter-frequency regularity, is unknown and to be estimated. In respect to the probabilistic point of vue, it is probabilized, and given a prior distribution $p(\rho)$.

3. Sequential Monte Carlo approach for global inversion

3.1. Rao-Blackwellised SMC algorithm

The posterior distribution $p(\mathbf{X}, \rho | \mathbf{Y})$ can be decomposed as: $p(\mathbf{X}, \rho | \mathbf{Y}) = p(\mathbf{X} | \rho, \mathbf{Y}) \cdot p(\rho | \mathbf{Y})$. Conditionally to ρ , the system is linear Gaussian: the conditional distributions $p(\mathbf{X} | \rho, \mathbf{Y})$ can be straightforwardly computed by Kalman filtering, including in this off-line context backward Kalman smoothing. On the other hand, the term $p(\rho | \mathbf{Y}) \propto p(\mathbf{Y} | \rho) \cdot p(\rho)$ can be evaluated (up to a normalising constant) for a given ρ using the likelihood term provided by the Kalman filter and prior distribution $p(\rho)$. Consequently, in order to exploit this conditional structure of the system, Kalman smoothers are applied and integrated in an interacting particle approach. This idea of mixing analytic integration (here Kalman evaluation of $p(\mathbf{X} | \rho, \mathbf{Y})$) with stochastic sampling is a variance reduction approach, known as Rao-Blackwellisation [3].

Similarly to [3], we choose to implement an efficient interacting particle approach, in order to estimate the marginal distribution $\eta(\rho) := p(\rho | \mathbf{Y}) \propto p(\mathbf{Y} | \rho) \cdot p(\rho)$. Sequential Monte Carlo (SMC) is a stochastic algorithm to sample from complex high-dimensional probability distributions. The principle (e.g., [4]) is to approximate a sequence of target probability distributions (η_n) by a large cloud of random samples termed particles ($(\zeta_n^k)_{1 \leq k \leq N_p} \in E^{N_p}$, E being called the state space. Between "times" $n-1$ and n , the particles evolve in state space E according to 2 steps:

- a **selection** step: every particle ζ_{n-1}^i is given a weight ω_i defined by a selection function g_n ($\omega_i = g_n(\zeta_{n-1}^i)$). By resampling (stochastic or deterministic), low-weighted particles vanish and are replaced by replicas of high-weighted ones.
- a **mutation** step: each selected particle $\widehat{\zeta}_{n-1}^i$ move, independently from the others, according to a Markov kernel M_n .

$$\left[\begin{array}{c} \zeta_{n-1}^1 \\ \vdots \\ \zeta_{n-1}^i \\ \vdots \\ \zeta_{n-1}^{N_p} \end{array} \right] \xrightarrow[\text{selection}]{g_n} \left[\begin{array}{ccc} \widehat{\zeta}_{n-1}^1 & \xrightarrow{M_n} & \zeta_n^1 \\ \vdots & & \vdots \\ \widehat{\zeta}_{n-1}^i & \xrightarrow{\quad} & \zeta_n^i \\ \vdots & & \vdots \\ \widehat{\zeta}_{n-1}^{N_p} & \xrightarrow{\quad} & \zeta_n^{N_p} \end{array} \right]$$

mutation

Evolving this way, the cloud of particles, and more precisely the occupation distribution $\eta_n^{N_p} = \frac{1}{N_p} \sum_{k=1}^{N_p} \delta_{\zeta_n^k}$ (sum of Dirac distributions), approximates for each n the theoretical distribution η_n defined recursively by the Feynman-Kac formulae, associated with the potentials g_n and kernels M_n [5].

Back to our objective of sampling from η , we then define the sequence of distributions $\eta_n(\rho)$:

$$\eta_n(\rho) \propto p(\mathbf{Y}|\rho)^{\alpha_n} \cdot p(\rho)$$

where $(\alpha_n)_{0 \leq n \leq n_f}$ is a sequence of number increasing from 0 to 1, so that: η_0 is prior distribution $p(\rho)$, easy to sample, η_{n_f} is target distribution η and sequence (η_n) admits a Feynman-Kac type structure with calculable selection functions g_n and Markov kernels chosen so that $\eta_n.M_n = M_n$ (Metropolis-Hastings for example). The distribution η is then interpreted as being the last distribution of a Feynman-Kac sequence, on which SMC can be performed, the estimator of η being the occupation distribution η^{N_p} extracted from the last cloud of particles.

3.2. Results

The occupation distribution η^{N_p} which approximates $\eta = p(\rho|\mathbf{Y})$ can be represented dimension by dimension via histograms (see figure 2).

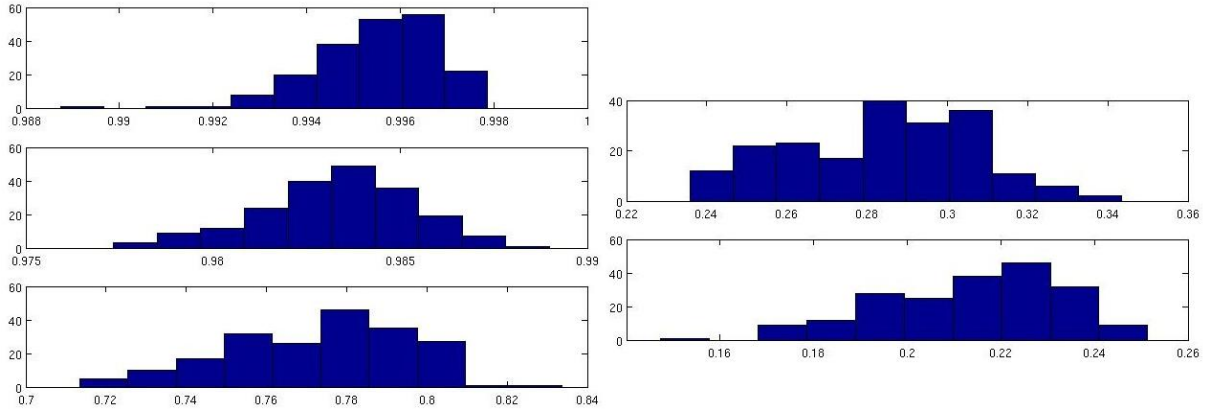


Figure 2. Estimation of a 5-dimensional ρ 's distribution with $N_p = 100$ particles.

For each frequency f_k , this approximation $\eta^{N_p} \simeq \eta$, associated with the theoretical conditioning relations

$$\begin{aligned} E(\mathbf{X}_k|\mathbf{Y}) &= E[E(\mathbf{X}_k|\rho, \mathbf{Y})|\mathbf{Y}] \\ Cov(\mathbf{X}_k|\mathbf{Y}) &= E[Cov(\mathbf{X}_k|\rho, \mathbf{Y})|\mathbf{Y}] + Cov[E(\mathbf{X}_k|\rho, \mathbf{Y})|\mathbf{Y}] \end{aligned}$$

can deliver estimators of respectively the mean and the covariance matrix of \mathbf{X}_k . Roughly speaking, the posterior estimation is performed by randomly picking a ρ_i from the final cloud of particles and computing associated samples of \mathbf{X}_k by a Kalman smoother conditionally to ρ_i . It is illustrated in figure 3, with a good agreement between the true state and estimated state.

Moreover, for any fixed area, the method provides estimators of the mean and marginal variance for every frequency, so that the results can be presented as frequential profiles, with marginal uncertainties (see figure 4). Even when the true (simulated) EM property profiles are chosen markedly divergent from the prior AR-type model, the method turns out to be robust. It results from the adaptive estimation of ρ which provides small values of ρ (i.e. weak correlation of EM properties for close frequencies) in the case of highly irregular true profiles.

4. Conclusion

An efficient statistical inference approach has been applied to estimate local material radioelectric properties from global EM scattering measurements. It combines intensive computations, meta-modeling and advanced sequential Monte Carlo techniques dedicated to frequency dynamic estimation.

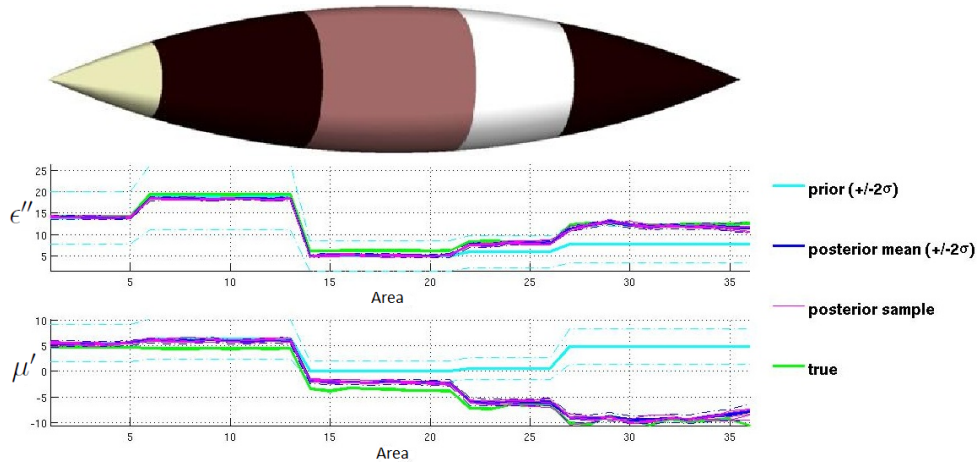


Figure 3. Estimated radioelectric properties of a $N_b = 5$ -block object ($N = 19$ areas), at fixed frequency f_k .

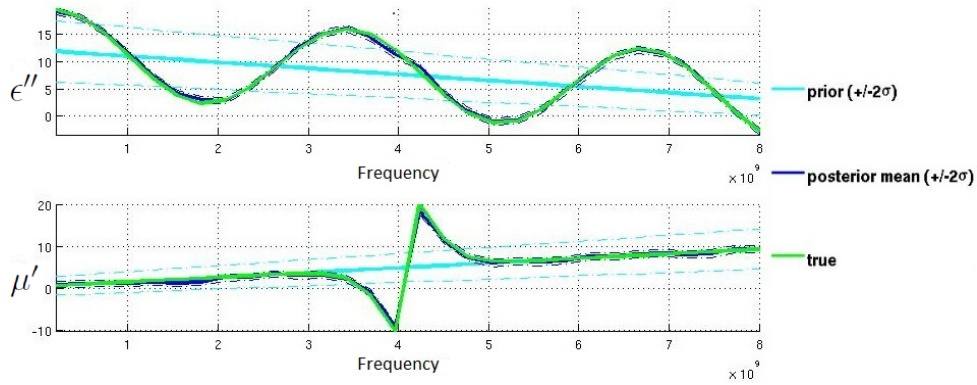


Figure 4. Estimated radioelectric properties ($K = 30$ frequencies) of a fixed area.

References

- [1] E. F. Knott, *Radar Cross Section Measurements*, SciTech Publishing, 2006.
- [2] S. M. Pandit, S-M Wu, *Time Series and System Analysis with Applications*, John Wiley & Sons, 1983.
- [3] J. S. Liu and R. Chen, *Sequential Monte Carlo for Dynamic Systems*, Journal of the American Statistical Association, 93, 1032-1044, 1998.
- [4] P. Del Moral, A. Doucet and A. Jasra, *Sequential Monte Carlo Methods for Bayesian Computation*, Bayesian Statistics 8, Oxford University Press, 2006.
- [5] P. Del Moral, *Feynman-Kac Formulae, Genealogical and Interacting Particle Approximations*, Springer New York (Series: Probability and Applications), 2004.